

---

# Genetic variants in glutamine metabolic pathway genes predict cutaneous melanoma-specific survival

Ka Chen<sup>1,2,3,#</sup>, Hongliang Liu<sup>2,3,#</sup>, Zhensheng Liu<sup>2,3</sup>, Wendy Bloomer<sup>2,3</sup>, Christopher I. Amos<sup>4</sup>,  
Jeffrey E. Lee<sup>5</sup>, Xin Li<sup>6</sup>, Hongmei Nan<sup>6</sup>, and Qingyi Wei<sup>2,3,7,\*</sup>

<sup>1</sup>Research Center for Nutrition and Food Safety, Institute of Military Preventive Medicine, Third Military Medical University, Chongqing 400038, PR China.

<sup>2</sup>Duke Cancer Institute, Duke University Medical Center, Durham, NC 27710, USA

<sup>3</sup>Department of Medicine, Duke University School of Medicine, Durham, NC 27710, USA

<sup>4</sup>Department of Community and Family Medicine, Geisel School of Medicine, Dartmouth College, Hanover, NH, 03755

<sup>5</sup>Department of Surgical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

<sup>6</sup>Department of Epidemiology, Richard M. Fairbanks School of Public Health, Indiana University, Indianapolis, IN 46202, USA

<sup>7</sup>Department of Population Health Sciences, Duke University School of Medicine, Durham, NC 27710, USA

# Co-first author

\*Correspondence author: Qingyi Wei, M.D., Ph.D., Duke Cancer Institute, Duke University Medical Center and Department of Population Health Sciences, Duke University School of Medicine, 905 S LaSalle Street, Durham, NC 27710, USA, Tel.: (919) 660-0562, E-mail: [qingyi.wei@duke.edu](mailto:qingyi.wei@duke.edu)

---

This is the author's manuscript of the article published in final edited form as:

Chen, K., Liu, H., Liu, Z., Bloomer, W., Amos, C. I., Lee, J. E., Li, X., Nan, H., & Wei, Q. (2019). Genetic variants in glutamine metabolic pathway genes predict cutaneous melanoma-specific survival. *Molecular Carcinogenesis*, 58(11), 2091–2103. <https://doi.org/10.1002/mc.23100>

---

**Funding Sources:** The MD Anderson Study was support by NIH/NCI R01 CA100264, 2P50CA093459 and R01CA133996 as well as by The University of Texas MD Anderson Cancer Center Various Donors Melanoma and Skin Cancers Priority Program Fund; the Miriam and Jim Mulva Research Fund; the McCarthy Skin Cancer Research Fund and the Marit Peterson Fund for Melanoma Research. The Harvard NHS/HPFS Study was in part supported by NIH/NCI R01 CA49449, P01 CA87969, UM1 CA186107 and UM1 CA167552. Ka Chen is sponsored by the China Scholarship Council for studying at the Duke University. Qingyi Wei was supported by start-up funds from Duke Cancer Institute, Duke University Medical Center and the P30 Cancer Center Support Grant (Grant ID: NIH CA014236).

**Running title:** Genetic variants in glutamine metabolic genes and CMSS

**Keywords:** Glutamine, cutaneous melanoma, genome-wide association study, single-nucleotide polymorphism, survival

**Abbreviations:** CM: cutaneous melanoma; SNP: single-nucleotide polymorphisms; GTEx, genotype tissue expression; GWAS: genome-wide association studies; MDACC: MD Anderson Cancer Center; CMSS: cutaneous melanoma-specific survival; NHS: Nurses' Health study; HPFS: Health Professionals; HWE, hardy Weinberg equilibrium; MAF, minor allele frequency; LD, linkage disequilibrium; HR: hazards ratio; CI: confidence interval; eQTL: expression quantitative trait loci; BFDP: bayesian false discovery probability; ROC: receiver operating characteristic; AUC: area under the curve.

**Conflict of interest statement:** None declared.

---

## Abstract

Glutamine dependence is a unique metabolic defect seen in cutaneous melanoma (CM), directly influencing the prognosis and treatment. Here, we investigated the associations between 6,025 common single-nucleotide polymorphisms (SNPs) in 77 glutamine metabolic pathway genes with CM-specific survival (CMSS) using two published genome-wide association study (GWAS) datasets. In the single-locus analysis, 76 SNPs were significantly associated with CMSS ( $P < 0.050$ , false-positive report probability  $< 0.2$  and Bayesian false discovery probability  $< 0.8$ ) in the discovery dataset, of which seven SNPs were replicated in the validation dataset and three SNPs (*HAL* rs17676826T>C, *LGSN* rs12663017T>A and *NOXRED1* rs8012548A>G) independently predicted CMSS, with an effect-allele attributed adjusted hazards ratio of 1.52 (95% confidence interval=1.19-1.93) and  $P < 0.001$ , 0.68 (0.54-0.87) and  $P = 0.002$  and 0.62 (0.46-0.83) and  $P = 0.002$ , respectively. The model including the number unfavorable genotypes (NUGs) of these three SNPs and clinical variables improved the five-year CMSS prediction ( $P = 0.012$ ). Further expression quantitative trait loci (eQTL) analysis found that the *LGSN* rs12663017 A allele was significantly associated with increased mRNA expression levels ( $P = 8.89 \times 10^{-11}$ ) in lymphoblastoid cell lines of the 1000 Genomes Project database; In the analysis of the genotype tissue expression (GTEx) project datasets, *HAL* rs17676826 C allele and *NOXRED1* rs8012548 G allele were significantly associated with their mRNA expression levels in sun-exposed skin of the lower leg ( $P = 6.62 \times 10^{-6}$  and  $1.37 \times 10^{-7}$ , respectively) and in sun-not-exposed suprapubic skin ( $P < 0.001$  and  $1.43 \times 10^{-8}$ , respectively). Taken together, these genetic variants of glutamine-metabolic pathway genes may be promising predictors of survival in CM patients.

---

## Introduction

Cutaneous melanoma (CM) is the most aggressive and treatment-resistant form of skin cancers (1). According to American Cancer Society, CM mortality remains high and stable in the past two decades, while dramatic declining trends are observed in most of other cancers (2, 3). Therefore, it is imperative to further identify the factors that predispose patients to poor survival outcome and to improve rationally prediction of prognosis and management of personalized treatment for CM patients.

Initiated by Warburg's seminal work, discoveries in metabolic alterations in cancer continue to raise tremendous interest (4-7). Glutamine dependency, the abnormally increased glutaminolysis, has been widely recognized as an important hallmark of cancer metabolism (5, 7-10). Glutamine, the most abundant free amino acid, plays a key role in cancer progression, including maintenance of carboxylic acid pools in the tricarboxylic acid cycle, sustaining cellular oxidative phosphorylation, and synthesis of the nonessential amino acids, purine, pyrimidines and fatty acids (9-11). To meet the dramatically increasing demand of glutamine in cancer, glutamine metabolism-related enzymes and transporters are overwhelmingly induced by oncogenes (7, 10, 12), and agents targeting glutamine metabolism including glutamine-mimetic compounds, inhibitors to glutaminase (GLS) or glutamate dehydrogenase (GDH) are proved to be effective in inhibiting cancer progression (10, 11, 13). Recently, glutamine addiction has also been identified in CM (1, 14, 15). Unlike glutamine-independence in melanocytes, melanoma cells, irrespective of their genetic backgrounds [e.g., mutated B-Raf Proto-Oncogene (BRAF), NRAS Proto-Oncogene (NRAS), or p53], all depend on glutamine for growth (1, 15-17),

---

consuming up to seven-fold more glutamine than melanocytes (15, 17, 18), whereby glutamine specifically provides a strong anaplerotic input, contributing to the biosynthesis of fatty acid and proline in hypoxia. In light of the glutamine's essential role in CM, it is important to comprehensively understand the regulation of glutamine metabolism in CM patients.

Multiple genetic alterations, either germline or somatic, have been reported to be involved in oncogenic transformation and progression of CM (19),(20, 21). Glutamine dependence in cancer may be due to one or a combination of deletions, polymorphisms or alterations in the genes of the glutamine/glutamine family amino acid metabolic process and glutamine transporters. Few studies have investigated the roles of genetic variants of the glutamine pathway genes in prognosis of melanoma. One study found that genetic variants of the glutamine transporter *SLC1A5* (*ASCT2*) were associated with prognosis of hepatocellular carcinoma patients (22). In the present study, we aimed to determine whether common genetic variants involved in the glutamine metabolism process are associated with CM-specific survival (CMSS) using two published genome-wide association study (GWAS) datasets, which may help identify promising prognostic biomarkers and support scientific foundation on the possible metabolism-based therapeutics.

## **Materials and Methods**

### ***Study populations and genotyping data***

Two GWAS datasets were used in the present study: the discovery dataset and the validation dataset. The discovery dataset was from The University of Texas MD Anderson Cancer Center

---

(MDACC) and the validation dataset was from the Nurses' Health study (NHS) and the Health Professionals Follow-up Study (HPFS) conducted by Harvard Brigham and Women's Hospital. The study protocols were approved by Institutional Review Boards at both MDACC and Harvard Brigham and Women's Hospital with a written consent from each of the subjects.

In the MDACC GWAS study, 858 non-Hispanic white patients diagnosed with histologically confirmed CM were accrued for a hospital-based case-control study between March 1998 and August 2008. The complete information (19, 21) for demographic and prospective clinicopathological data of the patients were obtained from a standard life-style questionnaire and/or extracted from patient medical charts. The genotypes were called by using the BeadStudio algorithm at John Hopkins University Center for Inherited Disease Research. Genome-wide imputation was conducted with the MACH software based on the 1000 Genomes Project, phase 1 v2 CEU data (23). SNPs with a minor allele frequency (MAF)  $\geq 0.05$ , a genotyping rate  $\geq 95\%$ , and Hardy-Weinberg equilibrium  $P$  values  $\geq 1 \times 10^{-5}$  were included in the final analysis. The MDACC dataset can be accessed at the Database of Genotypes and Phenotypes (dbGaP: <http://www.ncbi.nlm.nih.gov/gap>) with an accession number phs000187.v1.p1 (24). The detailed genotyping information and data quality control have been reported (19, 25).

In the Harvard NHS/HPFS GWAS study, the two cohorts of NHS and HPFS were established in 1976 and 1986, respectively (26). In the NHS, the information on CM development was first collected in 1984 and 90% of participants had completed health-related information for >20 years. In the HPFS, the related information was first collected in 1984 and the average follow-up rate

---

over 10 years was >90%. Eligible subjects in both cohorts were participants with histopathologically confirmed invasive melanoma, diagnosed at any time after baseline up to the 2008 follow-up cycle. All subjects were non-Hispanic whites in the United States. In the final analysis, 409 patients were included in the data after quality control. Genotyping was performed using the Illumina HumanHap610 array. Genome-Wide imputation was also performed using the MACH program based on the 1000 Genomes Project (Utah Residents with Northern and Western European Ancestry data, phase I v3) (27, 28). SNPs with imputation quality  $r^2 \geq 0.8$  and MAF  $\geq 0.05$  were included in the final analysis.

### ***Gene and SNP selection***

Based on the database of the Molecular Signatures Databases and literatures (13, 29, 30), 84 autosome genes of the glutamine/glutamine family amino acid metabolic process and glutamine transporters were selected to further investigation (**Supporting information Table 1**). After excluding seven genes in the X chromosome, SNPs within the remaining 77 genes and their 2-kb flanking regions were extracted from the MDACC GWAS dataset.

### ***In silico functional analysis as a biological validation***

For those validated SNPs as significant ones, bioinformatics functional prediction was performed firstly by using two online tools: RegulomeDB (<http://www.regulomedb.org>) and HaploReg (<http://archive.broadinstitute.org/mammals/haploreg/haploreg.php>) (31, 32). Then, the expression quantitation trait loci (eQTL) analysis was conducted by using data from multiple

---

sources: lymphoblastoid cell data of 373 European individuals from Genetic European Variation in Health and Disease Consortium (GEUVADIS) and the 1000 Genomes Project (phase I integrated release 3, March 2012) (27); the whole blood, skin or the subcutaneous adipose tissue data from the genotype tissue expression (GTEx) project (33); tumor and adjacent normal tissue data from the Cancer Genome Atlas (TCGA) database (34).

### ***Statistical methods***

CMSS was defined as the primary endpoint of the present study, for which survival time started at the date of diagnosis of CM and ended at the date of CM-related death or the last follow-up. Deaths with non-CM causes were considered censored. The associations between SNPs and CMSS (in an additive genetic model) were analyzed by both univariable and multivariable Cox proportional hazards regression models using the GenABEL package of R software, with adjustment for age, sex, Breslow thickness, tumor stage, tumor cell mitotic rate and ulceration of tumor in the MDACC GWAS dataset and for age and sex in the Harvard NHS/HPFS GWAS dataset. Bayesian measure of the false-positive report probability (FPRP) (35) correction was applied to limit the probability of false-positive findings as a relatively large number SNPs had been tested. We chose the less stringent FPRP for multiple test correction, because the vast majority (5089 out of 6209) of the SNPs included in the analysis were imputed with a high LD with the 1124 genotyped SNP (**Fig. 1A**). Only those SNPs with an FPRP value < 0.2 were considered worthy of subsequent validation in the Harvard NHS/HPFS GWAS dataset. We also used false discovery rate (FDR) and another Bayesian measure (BFDP: Bayesian false



---

discovery probability with a cutoff value of 0.8 for identifying noteworthy associations, which refines the criteria for FPRP) for identifying noteworthy associations (36). Linkage disequilibrium (LD) analysis was performed by using HaploView 4.2 according to European populations from the 1000 Genome Project with pairwise  $r^2=0.8$  as a cut-off value.

The stepwise Cox regression model including validated SNPs and clinical variables was performed to choose the independent SNPs. Meta-analysis was conducted using PLINK 1.09. Cochran's Q statistics and  $I^2$  were carried out to assess an inter-study heterogeneity. Fixed-effects models were used when no heterogeneity was found between two studies (Q test  $P$  values  $>0.1$ , and  $I^2 < 25.0\%$ ); otherwise, random-effects models were used. The number of unfavorable (risk) or protective genotypes (NUGs) was used as a genetic risk (protective) score to assess the combined effect of all independent and significant SNPs. Manhattan plot and quantile-quantile plot were performed using qqman package of R software. Kaplan-Meier survival curves and Log-rank tests were performed to visually evaluate the effects of NUGs on CMSS. The receiver operating characteristic (ROC) curve and time-dependent area under the curve (AUC) were constructed from the logistic regression model with the survival ROC package of R software. Statistical significance of the improvement in AUC was analyzed by the DeLong's test. All analyses were performed using SAS (version 9.1.4; SAS Institute, Cary, NC), unless otherwise specified. **Fig. 1A** provides the study follow chart, illustrating procedures of analyses performed in the present study.

## Results

---

### ***Basic characteristics of the two GWAS datasets***

The analysis included 858 patients from the MDACC GWAS dataset and 409 patients from the Harvard NHS/HPFS GWAS datasets, and basic characteristics of these subjects had been previously described (21) (**Supporting information Table 2**). In the MDACC GWAS dataset, the age of patients ranged between 17 and 94 years at diagnosis ( $52.4 \pm 14.4$  years), and there were more men (496, 57.8%) than women (362, 42.2%). Meanwhile, more patients had a stage I/II disease (709, 82.6%) than a stage III/IV disease (149, 17.4%), with a median follow-up time of 81.1 months, during which 95 (11.1%) patients died of CM. In addition, univariable Cox regression analysis indicated that age, sex, regional/distant metastasis, Breslow thickness, ulceration, and mitotic rate were significantly associated with CMSS. In the NHS/HPFS GWAS dataset, the age of the included patients was between 34 and 87 years at diagnosis ( $61.1 \pm 10.8$  years), and 66.3% (271) of the patients were women. Compared with that of the MDACC patients, the median follow-up time of the patients was relatively longer (179.0 months), during which 48 (11.5%) patients died of CM, and only age was significantly associated with CMSS in univariable Cox regression analysis of the NHS/HPFS GWAS dataset, because other clinicopathological variables were not available.

### ***SNPs and CMSS***

In the discovery MDACC GWAS dataset, Cox regression analysis was firstly performed to assess associations of a total of 6,209 common SNPs of the glutamine metabolism-related pathway genes with CMSS. Manhattan plot of associations between these variants and CMSS in

---

MDACC dataset is shown in **Fig. 1B**. The quantile-quantile plot of the observed  $P$  values showed a uniform distribution (**Fig. 1C**). In a single locus analysis, 252 SNPs were found to be significantly associated with CMSS at  $P < 0.05$  in an additive genetic model, of which 76 SNPs were still considered noteworthy after the correction by the FPRP  $< 0.2$  and BFDP  $< 0.8$ . Then, these 76 SNPs were further subjected to validation in the Harvard NHS/HPFS GWAS dataset. As shown in **Table 1**, seven SNPs in three genes identified in the discovery phase remained statistically significant ( $P < 0.05$ ), of which rs17676826 in histidine ammonia-lyase (*HAL*) was significantly associated with a poorer survival, while the other five SNPs of Lengsin (*LGSN*) and rs8012548 in NADP-dependent oxidoreductase domain-containing protein 1 (*NOXRED1*) were associated with a better survival in both datasets. Additionally, the noteworthy associations were also assessed by FDR (**Supporting information Table 3**). Subsequent meta-analysis of these SNPs from both datasets showed that these associations remained statistically significant, and there was no evidence for heterogeneity in these seven SNPs between the two GWAS datasets.

### ***Three independent SNPs as CM survival predictors***

We further performed functional prediction with RegulomeDB and HaploReg<sup>30</sup> for these validated SNPs (**Supporting information Table 4**). As indicated by RegulomeDB, the score of *HAL* rs17676826 was 4, and functional annotation of this SNP in HaploReg demonstrated that it overlaps with an enhancer, potentially disrupting four motifs [i.e., Zinc finger and BTB domain-containing protein 3 (*Zbtb3*)] and affecting the mRNA expression. As for the five SNPs in *LGSN*, LD analysis showed that they were in high LD ( $r^2 > 0.8$ ) (**Fig. 1D**), of which *LGSN*

---

rs12663017 may disrupt eight motifs [i.e., Forkhead Box A (FOXA), Histone Deacetylase 2 (HDAC2), and E1A Binding Protein P300 (EP300)] and affecting mRNA expression of the corresponding gene. Additionally, *LGSN* rs12663017 is located very close to the 3'-UTR (untranslated region, 247 bp at the 3' of *LGSN*), whereas SNPs rs2253428 and rs2253430 in complete LD ( $r^2=1$ ) with the lead variant rs12663017 were located at 3'-UTR. According to the HaploReg, *NOXRED1* rs8012548 may disrupt the motif of paired like homeodomain 2 (Pitx\_2) and spermatogenic leucine zipper 1 (Spz1), affecting the mRNA expression. Moreover, two highly linked SNPs with *NOXRED1* rs8012548 [i.e., transmembrane emp24 domain-containing protein 8 (*TMED8*) rs10141317,  $r^2 = 0.93$ ; *TMED8* rs3742737,  $r^2 = 0.94$ ] were identified as missense variants. The above-mentioned online functional predictions suggested that these SNPs were biologically functional.

Comprehensively considering the predicted functions,  $P$  values and LD, we selected three SNPs of *HAL* rs17676826, *LGSN* rs12663017 and *NOXRED1* rs8012548 as the tagSNPs. Then we used initial stepwise Cox regression analyses to identify whether these three SNPs were independent predictors of CMSS. The results suggested that these three tagSNPs were statistically significant independent predictors of CMSS (**Table 2**).

For each of the three independent SNPs, univariable and multivariable Cox regression analysis were further performed to evaluate their effects on death risk with adjustment of other clinicopathological covariates (**Table 3**). In the MDACC dataset, risk of death was significantly increased with the number of *HAL* rs17676826 C allele ( $P_{trend} = 0.001$ ) but was significantly decreased with the number of *LGSN* rs12663017 A and *NOXRED1* rs8012548 G alleles ( $P_{trend} =$

---

0.024 and 0.021, respectively). Similarly, consistent trends were observed in the Harvard NHS/HPFS dataset ( $P_{trend} = 0.029, 0.025$  and  $0.030$ , respectively) and in the MDACC and NHS/HPFS combined dataset ( $P_{trend} = 0.029, 0.004$  and  $0.030$ , respectively). In addition, regional association plots for variants in *HAL*, *LGSN* and *NOXRED1*, including the 50-kb regions flanking the neighborhoods of these genes, are shown in **Supporting information Fig. 1**.

### ***Combined effects of the three independent SNPs***

To better evaluate the joint effect of the three independent SNPs on risk of death, the risk genotypes (i.e., *HAL* rs17676826 TC+CC, *LGSN* rs12663017 TT and *NOXRED1* rs8012548 AA) were combined into one variable as a genetic score as the number of unfavorable genotypes (NUGs) (**Table 3**). The trend test indicated that an increased number of NUGs was associated with an increased risk of death in the MDACC ( $P < 0.001$ ), NHS/HPFS ( $P = 0.010$ ) and the combined ( $P = 0.012$ ) datasets. We further divided the combined NUGs into a low-risk group (0-1 NUGs) and a high-risk group (2-3 NUGs) and found that the HR for the high-risk group was 2.04 fold (CI = 1.32-3.15,  $P = 0.001$ ), 1.66 fold (CI = 0.94-2.92,  $P = 0.082$ ) and 1.90 fold (CI = 1.37-2.66,  $P < 0.001$ ) for the MDACC, NHS/HPFS and the combined datasets, respectively, compared with the low-risk group. For the illustrative purpose, Kaplan-Meier survival curves of these associations of the NUGs with CMSS are depicted in **Figs. 2A-2B and Supporting information Fig. 2**.

### ***Stratified analyses for associations between NUGs and CMSS***

---

Additional stratified analysis was carried out to investigate whether the combined effect of unfavorable genotypes on CMSS was modified by clinical variables in the MDACC dataset. Compared with those with 0-1 NUGs, individuals with 2-3 NUGs showed a poorer survival in the presence of clinicopathologic risk factors in the stratified subgroups of > 50 years old, male, female, with regional/distant metastasis, Breslow's thickness > 1mm, ulceration and mitotic rate > 1mm<sup>2</sup>, and no heterogeneity was observed among these subgroups (**Supporting information Table 4**).

#### ***The ROC curve and time dependent AUC***

We further estimated predictive value of the NUGs with time-dependent AUC and ROC curves using the combined MDACC and NHS/PFS datasets. As shown in **Fig. 2C**, the time-dependent AUC plot indicated an improved prediction performance with the addition of NUGs to the model with clinicopathologic risk factors (age and sex) from the beginning of the follow-up and remaining over time, compared with clinicopathologic factors only. As for classification of five-year CMSS (**Fig. 2D**), the AUCs were increased from 61.25% to 67.34% ( $P = 0.012$ ), which were statistically significant after adding NUGs into the model with clinicopathologic risk factors.

#### ***The eQTL analysis using 1000 Genomes Project database***

To evaluate correlations between SNPs and their corresponding mRNA expression levels, we primarily used the RNA-Seq data of lymphoblastoid cell lines from 373 European

---

descendants included in the database of the 1000 Genomes Project. As shown in **Figs. 3A**, *LGSN* rs12663017 TA and AA genotypes (or the A allele) were significantly associated with increased mRNA expression levels of *LGSN* (trend test in an additive model:  $P = 8.89 \times 10^{-11}$ ), but this trend was not observed for *HAL* rs17676826 ( $P = 0.341$ ) and *NOXRED1* rs8012548 ( $P = 0.079$ ).

### ***The eQTL analysis using the GTEx project database***

Then, we further conducted the eQTL analysis using data from the GTEx. In skin tissues from the donors, the *HAL* rs17676826C allele was associated with a significant increase in mRNA expression levels in sun-exposed skin of the lower leg ( $P = 6.62 \times 10^{-6}$ , **Fig. 3B-a**) and in sun-not-exposed suprapubic skin ( $P = 0.0011$ , **Fig. 3B-b**); the *NOXRED1* rs8012548 G allele was associated with a significant decrease in mRNA expression levels in sun-exposed skin of the lower leg ( $P = 1.37 \times 10^{-7}$ , **Fig. 3B-c**) and sun-not-exposed suprapubic skin ( $P = 1.43 \times 10^{-8}$ , **Fig. 3B-d**), and the two highly-linked variants (i.e., rs10141317 and rs3742737) were also associated with a significant decrease in *NOXRED1* mRNA expression levels in sun-exposed skin of the lower leg (rs10141317 T allele,  $P = 4.10 \times 10^{-7}$  and rs3742737 T allele,  $P = 3.10 \times 10^{-7}$ ) and sun-not-exposed suprapubic skin (rs10141317 T allele,  $P = 1.60 \times 10^{-8}$  and rs3742737 T allele,  $P = 1.10 \times 10^{-8}$ ); whereas there was no data about *LGSN* rs12663017 in skin tissues.

Additionally, according to the GTEx portal, *LGSN* the rs12663017A allele was associated with a significant increase in mRNA expression levels ( $P = 3.90 \times 10^{-9}$ , **Fig. 3B-e**) in the whole blood from the donors, and the same trends were observed in its six highly-linked variants (data

---

not shown); the *NOXRED1* rs8012548G allele was associated with a significant decrease in mRNA expression levels in subcutaneous adipose tissue ( $P = 2.20 \times 10^{-5}$ , **Fig. 3B-f**) from the donors.

### ***The eQTL analysis using the TCGA database***

We also performed SNPs and mRNA expression correlation analysis by using the expression data in tumor tissues from 473 cases of CM from the TCGA database; however, there was no significant association between *HAL* rs17676826 and its mRNA expression levels ( $P = 0.412$  in additive model); the other two tag SNPs were not included in the TCGA dataset.

Furthermore, associations between mRNA expression and lung cancer overall survival in TCGA database were investigated by using OncoLnc (<http://www.oncolnc.org/>). According to OncoLnc, CM patients with higher *NOXRED1* mRNA expression levels in tumor tissue showed a better overall survival ( $P = 0.025$ ) in TCGA database; Kaplan-Meier survival curves for *NOXRED1* in CM with the bottom quartile vs. top quartile of *NOXRED1* mRNA expression levels were shown in **Fig. 3C**. However, no significant associations were found between the *HAL* or *LGSN* mRNA expression and CM overall survival in TCGA database by using OncoLnc.

### **Discussion**

To our knowledge, this is the first report about the associations between genetic variants in the glutamine metabolism pathway genes and CMSS. Using publically available genotyping data from two published GWAS datasets, we revealed that genetic variants of *HAL* rs17676826, *LGSN*



---

rs12663017 and *NOXRED1* rs8012548 either individually or jointly modulated the survival of patients with CM. Remarkably, the effect was consistent across analyses of different datasets and through stratified analyses, and the genotype-survival association was pronounced even in the presence of clinicopathological risk factors, such as Breslow thickness and mitotic rate. Moreover, these genetic variants were found to influence their mRNA expression levels. These findings suggest that genetic variants in the glutamine metabolism pathway genes may have biological roles in CM progression, possibly through a mechanism of modulating expression of these genes, which would provide new scientific insights into metabolism-based therapeutics for cancer.

*HAL*, located on chromosome 12q23.1, encodes the histidine ammonia-lyase. Histidine is one of the glutamine family amino acids, which are disposed of through conversion to glutamate (37). *HAL*, mainly existing in the epidermis and liver, catabolizes histidine to *trans*-urocanic acid (UCA), an ultraviolet (UV) radiation-absorbing molecule in the stratum corneum, which can be photoisomerized to *cis*-UCA, when exposed to UV, especially UV-B (38-40). Moreover, *cis*-UCA mimics the effects of UV-B-mediated immuno-suppression, which has been recognized as an important factor related to skin cancer development (38-40). In the present study, we found that rs17676826, located in the intron region of *HAL*, was associated with a poorer survival in CM patients. The rs17676826 position overlaps with an enhancer activity cluster, which is classified as a genic enhancer by the 15-state core model and as a transcribed 3' enhancer by the 25-state model, according to the Haploreg database (32). Furthermore, histone modification markers H3K4me1, H3K4me4 and H3K27ac are all contributing to the chromatin state assignment at this

---

SNP location. Therefore, rs17676826 probably affects gene expression levels by modifying the accessibility of chromatin during the transcription. Consistent with that, the rs17676826 C allele was found to be associated with a significant increase in mRNA expression levels of *HAL* in the sun-exposed lower leg skin and sun-not-exposed suprapubic skin. Additional Haploreg data show that this SNP changes the match to some regulatory motifs, such as Zbtb3, which play important roles in cancer cell growth via gene expression of detoxification enzymes for reactive oxygen species (41). Overall, these provide a possible explanation to the mechanism underlying the observed association between rs17676826 and CMSS.

*LGSN*, located on chromosome 6q12, encodes a glutamine synthetase I family protein called lengsin that was previously reported to be a constitutive lens-specific protein but without the glutaminase activity (42). Recently, lengsin was identified as a novel tumor-associated antigen and revealed its essential role in cell survival (42, 43). In the present study, we found that CM patients with genotypes of *LGSN* rs12663017 TA+AA had a better CMSS. This SNP and its highly-linked SNPs, rs2253428 and rs2253430, are located very close to or within the 3'-UTR of the gene, which indicates that these SNPs may influence the fate of *LGSN* mRNA and thus proteosynthesis. Consistent with that, we found the *LGSN* rs12663017A allele was associated with a significant increase in mRNA expression levels of *LGSN* in whole blood and lymphoblastoid cell lines. Additionally, as indicated in the Haploreg database (32), rs12663017 can disrupt some important transcription regulators, including FOXA, EP300 and HDAC2, which suggests that this SNP is also likely affect gene expression by modifying the remodeling of chromatin during transcription.

---

*NOXRED1*, located on chromosome 6q12, encodes the NADP-dependent oxidoreductase domain-containing protein 1, formerly named C14orf148. According to the Go annotation (<http://www.geneontology.org/>), *NOXRED1* is a probable oxidoreductase and belongs to the pyrroline-5-carboxylate reductase-like protein (PYCR) family (44). In general, glutamine is degraded to glutamate, which then can consecutively be converted to proline by PYCR. Many studies have proved that proline is essential for cancer cell growth and that the elevated proline pool may enhance production of collagen and new extracellular matrix deposition, facilitating tumor invasion (18, 45-47). In the present study, our findings revealed that carriers of the *NOXRED1* rs8012548 G variant genotypes had a better CMSS. According to the annotation of HaploReg (32), the *NOXRED1* rs8012548G allele might have some effects on regulatory motifs, including Pitx2 and Spz1, which were associated with tumor progression in procollagen lysyl hydroxylase translation (48) and toll-like receptor (TLR) activation (49), respectively. Additionally, it should be noted that the association between rs8012548 and CMSS could have been correlated with the other two highly-linked missense variants (rs10141317 and rs3742737) in *TMED8*. These two *TMED8* variants may have substantial functions as they are located at the promoter enhancer and DNase I hypersensitive sites, likely to disrupt many motifs in various cells and tissues. Consistently, we found that the rs8012548 G allele, the rs10141317 T and rs3742737 T alleles were involved in transcriptional regulation as evidenced by eQTL analysis from the 1000 Genomes and GTEx projects. Notably, we found that these three SNPs were associated with significant decrease in mRNA expression levels in sun-exposed or sun-not-exposed skin, which supports an oncogenic effect of *NOXRED1*. However, in the TCGA

---

database, CM patients with higher *NOXRED1* mRNA expression levels in tumor tissue showed a better survival. These contradictions might be attributed to many factors including the limited samples of the two databases. Therefore, further functional investigation of *NOXRED1* are needed, especially for the experimental work.

There were some limitations in the present study. Firstly, there was no available information about glutamine or nutritional status, nutrition-based treatment and systemic therapies received by the patients. Secondly, although we adjusted in the models for variables (age, sex, Breslow thickness, regional/distant metastasis, ulceration, and mitotic rate) that could confound our observations of a genetic effect on CMSS for the discovery in the MDACC analysis, only age and sex were adjusted in the validation in the Harvard NHS/HPFS dataset. However, no heterogeneity was observed, when the two datasets were combined, which indicates that the observed effect of each SNP on CMSS from the two studies was consistent. Lastly, no direct biological experiments were conducted for functional validations. Further functional studies are warranted to investigate the exact function of these SNPs or genes on melanoma progression. Because we did not use the FDR as the multiple test correction, it is possible that our findings could be of false discovery, and thus additional validation is warranted. Also, additional larger validation studies with multiethnic groups are also needed to confirm our results, because our prognosis-predicting model was based on a non-Hispanic white patient population.

In conclusion, the present study identified the roles of genetic variants in *HAL* (rs17676826), *LGSN* (rs12663017) and *NOXRED1* (rs8012548) in CMSS as assessed in two independent GWAS datasets. Given the importance of glutamine metabolic alteration in the progression of

---

cancer cells, these genetic variants may represent promising prognostic biomarkers and potential subtype classification indicators for personalized metabolic therapies.

### **Supplementary Material**

Supplementary material can be found online.

### **Acknowledgments**

We would like to thank the John Hopkins University Center for Inherited Disease Research for conducting high-throughput genotyping for this study. We thank the participants and staff of the Nurses' Health Study (NHS) and Health Professionals Follow-up Study (HPFS) for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. We also thank TCGA research network. We deeply appreciate all the individuals who participated in this project. The authors assume full responsibility for analyses and interpretation of these data.

### **Authors' contributions**

The authors' responsibilities were as follows—KC: was the manuscript principal author and was responsible for the conception and design of the project, data analysis and manuscript writing; HL: was responsible for design the project, data analysis and contributed to manuscript writing; ZL: helped design the project and revise the final manuscript; WB, DZ, CA, SF, JEL, XL

---

and HN: contributed to data collection and preparation of the manuscript; and QW (Principal Investigator): was responsible for the conception and design of the project and financial support and contributed to the preparation and revision of the manuscript. All authors reviewed and approved the final version of the manuscript.

### **Data Availability Statement**

The MDACC dataset can be accessed at the Database of Genotypes and Phenotypes (dbGaP: <http://www.ncbi.nlm.nih.gov/gap>) with an accession number phs000187.v1.p1. The data using in the expression quantitation trait loci analysis were from the Genetic European Variation in Health and Disease Consortium and the 1000 Genomes Project (phase I integrated release 3, March 2012), the genotype tissue expression project and the Cancer Genome Atlas database.

### **References**

1. Ratnikov BI, Scott DA, Osterman AL, Smith JW, Ronai ZA. Metabolic rewiring in melanoma. *Oncogene* 2017; 36: 147-157.
2. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. *CA Cancer J Clin* 2017; 67: 7-30.
3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016; 66: 7-30.
4. Corbet C, Feron O. Cancer cell metabolism and mitochondria: Nutrient plasticity for TCA cycle fueling. *Biochim Biophys Acta* 2017; 1868: 7-15.
5. Pavlova NN, Thompson CB. The Emerging Hallmarks of Cancer Metabolism. *Cell Metab* 2016; 23: 27-47.

- 
6. Coller HA. Is cancer a metabolic disease? *Am J Pathol* 2014; 184: 4-17.
  7. Cantor JR, Sabatini DM. Cancer cell metabolism: one hallmark, many faces. *Cancer Discov* 2012; 2: 881-898.
  8. Martinez-Outschoorn UE, Peiris-Pages M, Pestell RG, Sotgia F, Lisanti MP. Cancer metabolism: a therapeutic perspective. *Nat Rev Clin Oncol* 2017; 14: 11-31.
  9. Hensley CT, Wasti AT, DeBerardinis RJ. Glutamine and cancer: cell biology, physiology, and clinical opportunities. *J Clin Invest* 2013; 123: 3678-3684.
  10. Zhang J, Pavlova NN, Thompson CB. Cancer cell metabolism: the essential role of the nonessential amino acid, glutamine. *EMBO J* 2017; 36: 1302-1315.
  11. Altman BJ, Stine ZE, Dang CV. From Krebs to clinic: glutamine metabolism to cancer therapy. *Nat Rev Cancer* 2016; 16: 619-634.
  12. Yang L, Venneti S, Nagrath D. Glutaminolysis: A Hallmark of Cancer Metabolism. *Annu Rev Biomed Eng* 2017; 19: 163-194.
  13. Bhutia YD, Babu E, Ramachandran S, Ganapathy V. Amino Acid transporters in cancer and their relevance to "glutamine addiction": novel targets for the design of a new class of anticancer drugs. *Cancer Res* 2015; 75: 1782-1788.
  14. Pan M, Reid MA, Lowman XH, et al. Regional glutamine deficiency in tumours promotes dedifferentiation through inhibition of histone demethylation. *Nat Cell Biol* 2016; 18: 1090-1101.
  15. Scott DA, Richardson AD, Filipp FV, et al. Comparative metabolic flux profiling of melanoma cell lines: beyond the Warburg effect. *J Biol Chem* 2011; 286: 42626-42634.

- 
16. Hernandez-Davies JE, Tran TQ, Reid MA, et al. Vemurafenib resistance reprograms melanoma cells towards glutamine dependence. *J Transl Med* 2015; 13: 210.
  17. Ratnikov B AP, Ronai ZA, Smith JW, Osterman AL, Scott DA. Glutamate and asparagine cataplerosis underlie glutamine addiction in melanoma. *Oncotarget* 2015; 6: 7379-7389.
  18. Filipp FV, Ratnikov B, De Ingeniis J, Smith JW, Osterman AL, Scott DA. Glutamine-fueled mitochondrial metabolism is decoupled from glycolysis in melanoma. *Pigment cell & melanoma research* 2012; 25: 732-739; e-pub ahead of print 2012/08/01.
  19. Amos CI, Wang LE, Lee JE, et al. Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. *Hum Mol Genet* 2011; 20: 5012-5023.
  20. Liu H, Liu Z, Wang Y, et al. Functional variants in DCAF4 associated with lung cancer risk in European populations. *Carcinogenesis* 2017; 38: 541-551.
  21. Liu S, W Y, Xue W, et al. Genetic variants in the genes encoding rho GTPases and related regulators predict cutaneous melanoma-specific survival. *Int J Cancer* 2017; 141: 721-730.
  22. Ge NJ, Shi ZY, Yu XH, et al. Genetic Variants in ASCT2 Gene are Associated with the Prognosis of Transarterial Chemoembolisation-Treated Early-Stage Hepatocellular Carcinoma. *Asian Pac J Cancer Prev.* 2015;16(9):4103-7.
  23. Li Y, Willer CJ, Ding J, Scheet P, Abecasis GR. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet Epidemiol* 2010; 34: 816-834.
  24. Mailman MD FM, Jin Y, Kimura M, et al. The NCBI dbGaP database of genotypes and phenotypes. *Nat Genet* 2007; 39: 1181-1186.



- 
25. Simard EP WE, Siegel R, Jemal A. Cancers with increasing incidence trends in the United States 1999 through 2008. *CA Cancer J Clin* 2012; 62: 118-128.
  26. Song F, Amos CI, Lee JE, et al. Identification of a melanoma susceptibility locus and somatic mutation in TET2. *Carcinogenesis* 2014; 35: 2097-2101.
  27. Lappalainen T, Sammeth M, Friedlander MR, et al. Transcriptome and genome sequencing uncovers functional variation in humans. *Nature* 2013; 501: 506-511.
  28. Li X, Liang L, Zhang M, et al. Obesity-related genetic variants, human pigmentation, and risk of melanoma. *Human genetics* 2013; 132: 793-801.
  29. Scalise M, Pochini L, Galluccio M, Indiveri C. Glutamine transport. From energy supply to sensing and beyond. *Biochim Biophys Acta* 2016; 1857: 1147-1157.
  30. Pochini L, Scalise M, Galluccio M, Indiveri C. Membrane transporters for the special amino acid glutamine: structure/function relationships and relevance to human health. *Front Chem* 2014; 2: 61.
  31. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Res* 2012; 40: D930-934.
  32. Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res* 2016; 44: D877-881.
  33. genomics. GCH. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 2015; 348: 648-660.

- 
34. Cancer Genome Atlas N. Genomic Classification of Cutaneous Melanoma. *Cell* 2015; 161(7): 1681-1696; doi 10.1016/j.cell.2015.05.044.
35. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. *J Natl Cancer Inst.* 2004;96(6):434-42.
36. Wakefield J.A Bayesian measure of the probability of false discovery in genetic epidemiology studies. *Am J Hum Genet.* 2007;81 (2):208-27.
37. Daye D, Wellen KE. Metabolic reprogramming in cancer: unraveling the role of glutamine in tumorigenesis. *Semin Cell Dev Biol* 2012; 23: 362-369.
38. Welsh MM, Karagas MR, Applebaum KM, Spencer SK, Perry AE, Nelson HH. A role for ultraviolet radiation immunosuppression in non-melanoma skin cancer as evidenced by gene-environment interactions. *Carcinogenesis* 2008; 29: 1950-1954.
39. IR. S. Factors controlling the expressed activity of histidine ammonia-lyase in the epidermis and the resulting accumulation of urocanic acid. *Biochem J* 1981; 194: 829-838.
40. Walterscheid JP, Nghiem DX, Kazimi N, et al. Cis-urocanic acid, a sunlight-induced immunosuppressive factor, activates immune suppression via the 5-HT<sub>2A</sub> receptor. *Proc Natl Acad Sci U S A* 2006; 103: 17420-17425.
41. Lim J-H. Zinc finger and BTB domain-containing protein 3 is essential for the growth of cancer cells. *BMB Reports* 2014; 47: 405-410.
42. Nakatsugawa M, Hirohashi Y, Torigoe T, et al. Novel spliced form of a lens protein as a novel lung cancer antigen, Lentsin splicing variant 4. *Cancer Sci* 2009; 100: 1485-1493.

- 
43. Nakatsugawa M, Horie K, Yoshikawa T, et al. Identification of an HLA-A\*0201-restricted cytotoxic T lymphocyte epitope from the lung carcinoma antigen, Lentsin. *Int J Oncol* 2011; 39: 1041-1049.
  44. Watford. M. Glutamine metabolism and function in relation to proline synthesis and the safety of glutamine and proline supplementation. *J Nutr* 2008; 138: 2003S-2007S.
  45. Phang JM, Liu W, Hancock CN, Fischer JW. Proline metabolism and cancer: emerging links to glutamine and collagen. *Curr Opin Clin Nutr Metab Care* 2015; 18: 71-77.
  46. Phang JM, Liu W, Hancock C, Christian KJ. The proline regulatory axis and cancer. *Front Oncol* 2012; 2: 60.
  47. Ding J KM, Su L, Xue L, et al. Human mitochondrial pyrroline-5-carboxylate reductase 1 promotes invasiveness and impacts survival in breast cancers. *Carcinogenesis* 2017; 38: 519-531.
  48. Hjalte TA AB, Murray JC. PITX2 regulates procollagen lysyl hydroxylase (PLOD) gene expression: implications for the pathology of Riegersyndrome. . *J Cell Biol* 2001; 152: 545-552.
  49. Wang Y ZS. Evolutionary and functional epitopes of the Spätzle protein New insights into activation of the Toll receptor. *Cell Mol Life Sci* 2009; 66: 1595-1602.

---

## Figure Legends

**Figure 1.** Screening for independent functional SNPs in the glutamine metabolic pathway genes to predict CMSS. (A) Study flowchart; (B) Manhattan plot and (C) quantile-quantile plot of associations between variants of the glutamine metabolic pathway and CMSS in MDACC dataset. There are 252 SNPs with  $P < 0.05$  and 76 SNPs with  $FPRP < 0.2$  and  $BFDP < 0.8$  in the total 6,209 SNPs of glutamine metabolic pathway. The blue horizontal line indicates  $P = 0.05$ . The red horizontal line indicates  $FPRP = 0.2$  and  $BFDP = 0.8$ . (D) LD plots of validated five SNPs in *LGSN*. CMSS, cutaneous melanoma-specific survival; MDACC, The University of Texas MD Anderson Cancer Center; SNP, single-nucleotide polymorphism; BFDP, bayesian false discovery probability; GWAS, genome-wide association study; HWE, hardy Weinberg equilibrium; MAF, minor allele frequency; LD, linkage disequilibrium.

**Figure 2.** The combined risk genotypes and survival prediction. Kaplan-Meier survival curves of the combined risk genotypes in the MDACC + NHS/HPFS datasets (A), and dichotomized groups of the NUG in the MDACC + NHS/HPFS datasets (B); ROC curves and AUC estimation for prediction of melanoma-specific survival using the MDACC + NHS/HPFS datasets, (C) Time-dependent AUC estimation, based on age, sex and the risk genotypes of the three independent genes, and (D) Five-year melanoma-specific survival ROC curves ( $P = 0.012$ ). NUG, number of unfavorable genotypes; MDACC, The University of Texas MD Anderson Cancer Center. ROC, Receiver operating characteristic; AUC, Time-dependent area under the ROC curve.

**Figure 3.** Associations between the risk genotypes and their corresponding mRNA expression levels. (A) The eQTL in 373 Europeans from the 1000 Genomes Project for (a) *HAL* rs17676826, (b) *LGSN* rs12663017 and (c) *NOXRED1* rs8012548 in the additive model; (B) the eQTL from the GTEx project using the additive model for *HAL* rs17676826 in skin tissues from sun-exposed lower leg skin (a) and sun-not-exposed suprapubic (b), *NOXRED1* rs8012548 in skin tissues from sun-exposed lower leg (c) and sun-not-exposed suprapubic

---

(d), *LGSN* rs12663017 in the whole blood (e), and *NOXRED1* rs8012548 in subcutaneous adipose (e). (C) Kaplan-Meier survival plot for *NOXRED1* in CM patients from TCGA database. Survival analyses based on different mRNA expression levels of *NOXRED1* in the bottom quartile vs. top quartile in 458 CM patients of European descents from TCGA database, which were analyzed by OncoLnc (<http://www.oncolnc.org>). eQTL, expression quantitative trait loci analysis. GTEx, Genotype-Tissue Expression; CM, cutaneous melanoma; TCGA, The Cancer Genome Atlas.

**Table 1.** Meta-analysis of seven validated SNPs of genes in the glutamine metabolic pathway using two published MDACC and NHS/HPFS melanoma GWAS datasets

SNP	Allele <sup>1</sup>	Gene	Position	MDACC (n=858)					NHS/HPFS (n=409)			Meta-analysis			
				EAf	HR (95% CI) <sup>2</sup>	P <sup>2</sup>	FPRP	BFDP <sup>3</sup>	EAf	HR (95% CI) <sup>4</sup>	P <sup>4</sup>	P <sub>het</sub> <sup>5</sup>	I <sup>2</sup>	HR (95% CI) <sup>6</sup>	P <sup>6</sup>
rs17676826 <sup>#</sup>	T/C	HAL	12q23.1	0.31	1.48 (1.1-1.99)	0.010	0.081	0.570	0.28	1.59 (1.05-2.42)	0.029	0.784	0	1.52 (1.19-1.93)	7.46x 10 <sup>-4</sup>
rs1723518 <sup>\$</sup>	G/A	LGSN	6q12	0.45	0.71 (0.53-0.95)	0.022	0.165	0.292	0.48	0.63 (0.41-0.94)	0.025	0.645	0	0.68 (0.54-0.87)	1.80 x 10 <sup>-3</sup>
rs9341780 <sup>\$</sup>	G/T	LGSN	6q12	0.45	0.71 (0.53-0.95)	0.022	0.165	0.292	0.48	0.63 (0.41-0.94)	0.025	0.644	0	0.68 (0.54-0.87)	1.80 x 10 <sup>-3</sup>
rs1007519 <sup>\$</sup>	C/T	LGSN	6q12	0.45	0.71 (0.53-0.96)	0.024	0.179	0.212	0.48	0.62 (0.41-0.94)	0.025	0.600	0	0.68 (0.53-0.86)	1.53 x 10 <sup>-3</sup>
rs1711934 <sup>\$</sup>	G/C	LGSN	6q12	0.45	0.71 (0.53-0.96)	0.024	0.179	0.253	0.48	0.63 (0.42-0.96)	0.032	0.687	0	0.69 (0.54-0.87)	2.02 x 10 <sup>-3</sup>
rs12663017 <sup>\$</sup>	T/A	LGSN	6q12	0.45	0.71 (0.53-0.96)	0.024	0.179	0.738	0.48	0.64 (0.41-0.94)	0.025	0.644	0	0.68 (0.54-0.87)	1.79x 10 <sup>-3</sup>
rs8012548 <sup>\$</sup>	A/G	NOXRED1	14q24.3	0.24	0.65 (0.45-0.94)	0.021	0.169	0.713	0.28	0.56 (0.33-0.94)	0.029	0.645	0	0.62 (0.46-0.83)	1.57x 10 <sup>-3</sup>

Abbreviations: SNP, single nucleotide polymorphism; GWAS, genome-wide association study; MDACC, The University of Texas M.D. Anderson Cancer Center; NHS/HPFS, Nurses' Health Study/Health Professionals Follow-up Study; EAF, effect allele frequency; HR, hazards ratio; CI, confidence interval; FPRP, false-positive report probability; BFDP, Bayesian false discovery probability.

P<sub>het</sub>, P value for heterogeneity by Cochrane's Q test;

<sup>1</sup>Reference allele/effect allele;

<sup>2</sup>Adjusted for age, sex, Breslow thickness, distant/regional metastasis, ulceration and mitotic rate in additive model;

<sup>3</sup>With a prior probability of 0.10 to detect an HR of 2.0 for an association with variant genotypes or minor alleles of the SNPs with P < 0.05.

<sup>4</sup>Adjusted for age and sex in additive model;

<sup>5</sup>P<sub>het</sub>, P value for heterogeneity by Cochrane's Q test;

<sup>6</sup>Meta-analysis in the fix-effect model;

<sup>#</sup> Genotyped SNP in the MDACC study;

<sup>\$</sup> Imputed SNP in the MDACC study.

**Table 2.** Independent predictors of CMSS as obtained from the stepwise Cox regression analysis of selected variables from the MDACC dataset

Parameter <sup>1</sup>	Category <sup>2</sup>	Frequency	HR (95% CI)	P
Age	≤50/>50	371/487	1.02 (1.01-1.04)	0.049
Sex	Female/Male	362/496	1.45 (0.91-2.32)	0.116
Regional/distant metastasis	No/Yes	709/149	4.48 (2.89-6.96)	<.0001
Breslow thickness(mm)	≤1/>1	347/511	1.15 (1.09-1.21)	<.0001
Ulceration	No/Yes	681/155	2.86 (1.86-4.41)	<.0001
Mitotic rate (mm <sup>2</sup> )	≤1/>1	275/583	2.48 (1.21-5.09)	0.013
<i>HAL</i> rs17676826 T>C	TT/TC/CC	410/361/87	1.46 (1.08-1.97)	0.015
<i>LGSN</i> rs12663017T>A	TT/TA/AA	271/400/187	0.71 (0.52-0.96)	0.025
<i>NOXRED1</i> rs8012548 A>G	AA/AG/GG	485/327/46	0.67 (0.47-0.97)	0.032

Abbreviations: CMSS, cutaneous melanoma-specific survival; MDACC, The University of Texas MD Anderson Cancer Center; HR, hazards ratio; CI, confidence interval;

<sup>1</sup> Stepwise analysis included age, sex, regional/distant metastasis, Breslow thickness, ulceration, mitotic rate and three SNPs (rs17676826 in *HAL*, rs12663017 in *LGSN* and rs8012548 in *NOXRED1*) in three genes.

<sup>2</sup> The “category” was used as the reference.

**Table 3.** Associations between three independent SNPs in the glutamine metabolic pathway genes and CMSS of patients in the MDACC study and NHS/HPFS study

Genotype	MDACC (n=858)				NHS/HPFS (n=409)				MDACC + NHS/HPFS (n=1267)			
	Frequency		Multivariate analysis <sup>1</sup>		Frequency		Multivariate analysis <sup>2</sup>		Frequency		Multivariate analysis <sup>3</sup>	
	All	Death (%)	HR (95%CI)	P	All	Death (%)	HR (95%CI)	P	All	Death (%)	HR (95%CI)	P
<b>HAL rs17676826 T&gt;C</b>												
TT	410	36 (8.78)	1.00		214	21 (9.81)	1.00		624	57 (9.13)	1.00	
TC	361	50 (13.85)	2.19 (1.39-3.45)	<0.001	161	19 (11.80)	1.26 (0.68-2.34)	0.467	522	69 (13.22)	1.60 (1.12-2.27)	0.009
CC	87	9 (10.34)	1.61 (0.76-3.40)	0.212	34	8 (23.53)	2.93 (1.29-6.64)	0.010	121	17 (14.05)	1.83 (1.06-3.15)	0.030
Trend test				0.010				0.029				0.004
TC+CC	448	59 (13.17)	2.07 (1.34-3.21)	0.001	195	27 (13.85)	1.51 (0.85-2.68)	0.156	643	86 (13.37)	1.64 (1.17-2.29)	0.004
<b>LGSN rs12663017 T&gt;A</b>												
TT	271	38 (14.02)	1.00		109	16 (14.68)	1.00		380	54 (14.21)	1.00	
TA	400	45 (11.25)	0.88 (0.56-1.37)	0.559	207	28 (13.53)	0.92 (0.50-1.71)	0.803	607	73 (12.03)	0.80 (0.56-1.14)	0.218
AA	187	12 (6.42)	0.44 (0.22-0.87)	0.018	93	4 (4.30)	0.27 (0.09-0.81)	0.019	280	16 (5.71)	0.36 (0.21-0.64)	<0.001
Trend test				0.024				0.025				0.0004
TA+AA	587	57 (9.71)	0.73 (0.48-1.12)	0.153	300	32 (10.67)	0.71 (0.39-1.29)	0.261	887	89 (10.03)	0.66 (0.47-0.92)	0.016
<b>NOXRED1 rs8012548 A&gt;G</b>												
AA	485	59 (12.16)	1.00		215	32 (14.88)	1.00		700	91 (13.00)	1.00	0.028
AG	327	34 (10.40)	0.79 (0.51-1.21)	0.274	162	15 (9.26)	0.62 (0.34-1.15)	0.13	489	49 (10.02)	0.78 (0.55-1.10)	0.152
GG	46	2 (4.35)	0.21 (0.05-0.86)	0.030	32	1 (3.13)	0.20 (0.03-1.46)	0.112	78	3 (3.85)	0.27 (0.09-0.86)	0.027
Trend test				0.021				0.030				0.012
AG+GG	373	36 (9.65)	0.68 (0.45-1.05)	0.081	194	16 (8.25)	0.55 (0.30-1.00)	0.051	567	52 (9.17)	0.70 (0.50-0.99)	0.041
<b>Number of risk genotypes<sup>4</sup></b>												
0	134	9 (6.72)	1.00		80	4 (5.00)	1.00		214	13 (6.07)	1.00	
1	319	27 (8.46)	1.92 (0.89-4.13)	0.097	164	19 (11.59)	2.50 (0.85-7.35)	0.096	483	46 (9.52)	1.64 (0.89-3.04)	0.115
2	330	48 (14.55)	3.03 (1.47-6.27)	0.003	140	19 (13.57)	2.91 (0.99-8.56)	0.053	470	67 (14.26)	2.62 (1.44-4.74)	0.002
3	75	11 (14.67)	3.98 (1.59-9.98)	0.003	25	6 (24.00)	5.52 (1.56-19.56)	0.008	100	17 (17.00)	3.39 (1.64-6.99)	0.001
Trend test				<0.001				0.010				< 0.001
0-1	453	36 (7.95)	1.00		244	23 (9.43)	1.00		697	59 (8.46)	1.00	
2-3	405	59 (14.57)	2.04 (1.32-3.15)	0.001	165	25 (15.15)	1.66 (0.94-2.92)	0.082	570	84 (14.74)	1.90 (1.37-2.66)	<0.001

Abbreviations: SNP, single nucleotide polymorphisms; CMSS, cutaneous melanoma-specific survival; GWAS, genome-wide association study; MDACC, The University of Texas MD Anderson Cancer Center; NHS/HPF, Nurses' Health Study/Health Professionals Follow-up Study; HR, hazards ratio; CI, confidence interval;

<sup>1</sup>Adjusted for age, sex, Breslow thickness, distant/regional metastasis, ulceration and mitotic rate in Cox models of SNPs and CMSS in MDACC study;

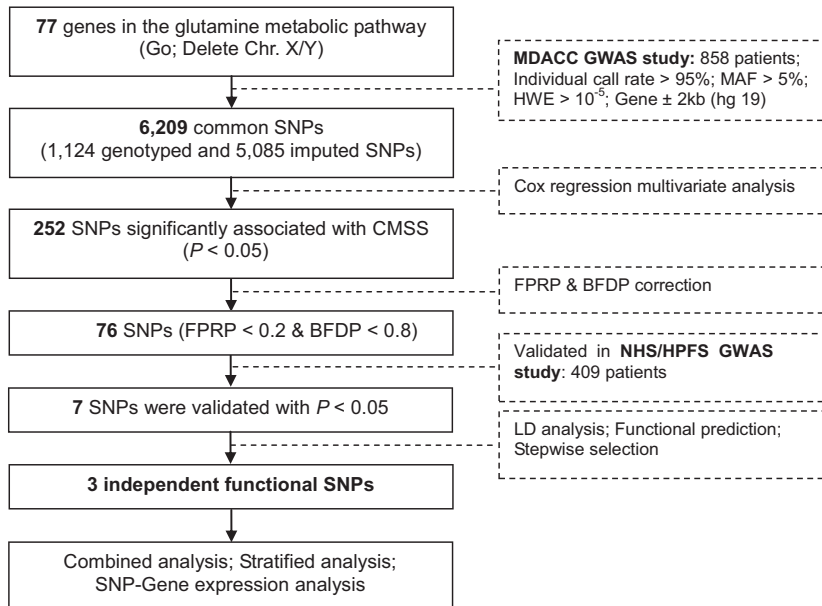
<sup>2</sup>Adjusted for age and sex in NHS/HPFS study;

<sup>3</sup>Adjusted for age and sex in MDACC and NHS/HPFS study;

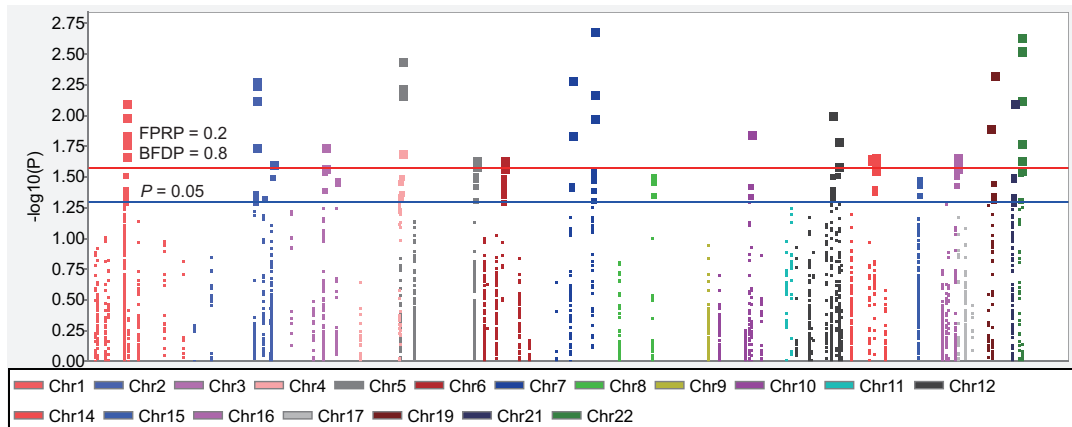
<sup>4</sup>Risk genotypes include *HAL* rs17676826 TC+CC, *LGSN* rs12663017 TT and *NOXRED1* rs8012548 AA.



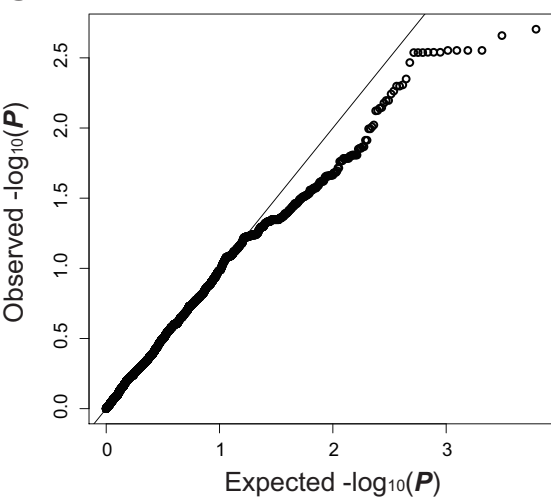
A



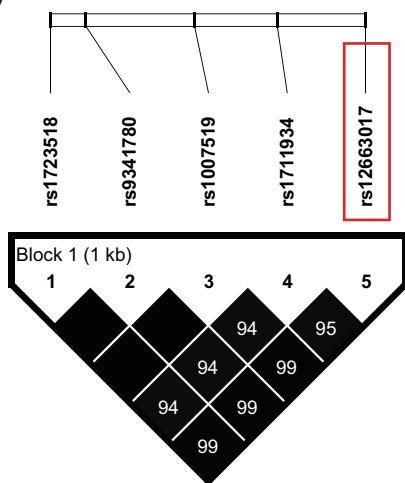
B

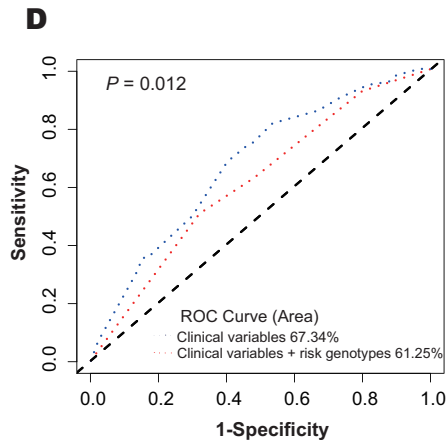
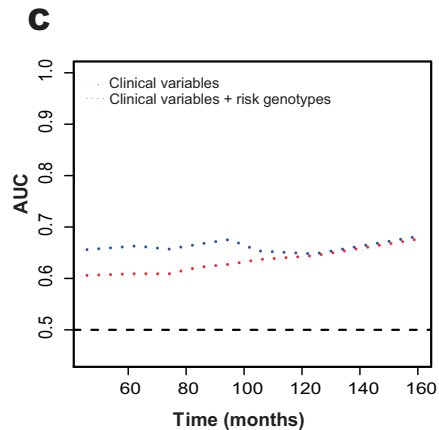
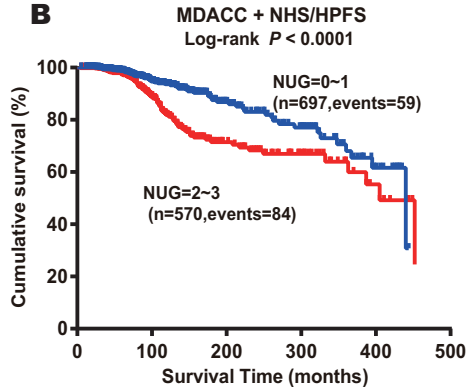
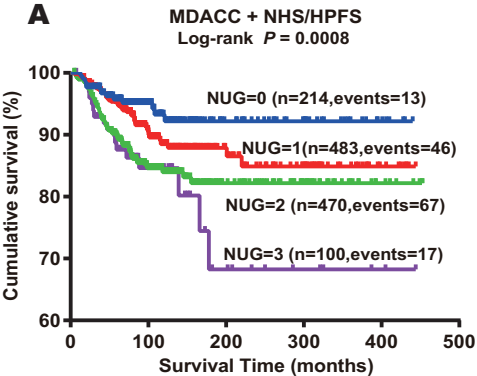


C

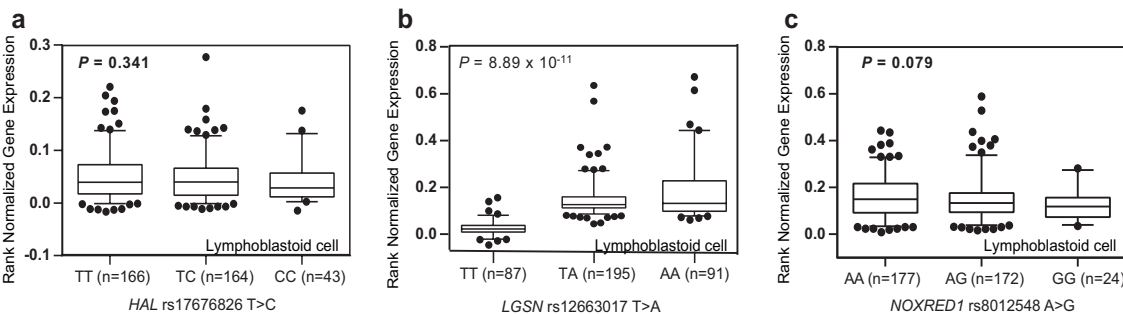


D

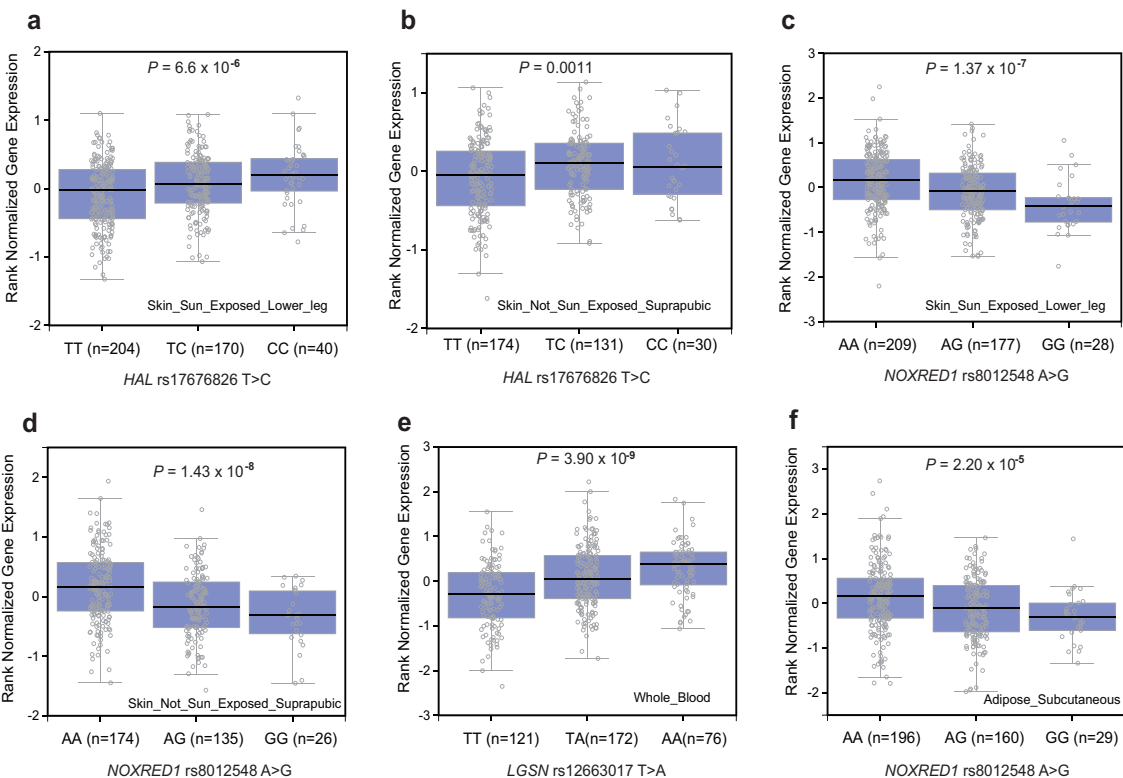




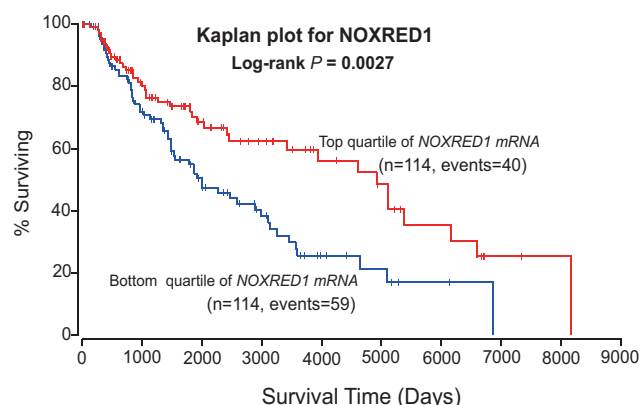
## A. The eQTL analysis in 1000 Genomes Project



## B. The eQTL analysis in GTEx Project



## C. Kaplan-Meier survival plot in TCGA database



---

## Supporting information Figure Legends

**Supporting information Figure 1.** Regional association plots for three independent SNPs in MDACC study. SNPs in the region of 50 kb up or downstream of *HAL* rs17676826 (A), *LGSN* rs12663017 (B) and *NOXRED1* rs8012548 (C) SNPs, Single nucleotide polymorphisms; MDACC, The University of Texas MD Anderson Cancer Center.

**Supporting information Figure 2.** The combined risk genotypes and survival prediction. Kaplan-Meier survival curves of the combined risk genotypes in the MDACC dataset (A) and the NHS/HPFS dataset (C), and dichotomized groups of the NUG in the MDACC dataset (B), the NHS/HPFS dataset (D). NUG, number of unfavorable genotypes; MDACC, The University of Texas MD Anderson Cancer Center.

**Supporting information Table 1.** List of 84 selected genes in the glutamine metabolic-related pathway genes

Dataset	Name of pathway	Number of genes
GO	CARBON NITROGEN LIGASE ACTIVITY WITH GLUTAMINE AS AMIDO N DONOR	12
GO	GLUTAMINE FAMILY AMINO ACID BIOSYNTHETIC PROCESS	20
GO	GLUTAMINEFAMILYAMINOACIDCATABOLIC PROCESS	24
GO	GLUTAMINEFAMILYAMINOACIDMETABOLIC PROCESS	65
GO	GLUTAMINE FAMILY AMINO ACID METABOLIC PRIC PROCESS	14
Literature*(1)	GLUTAMINE TRANSPORTERS	12
<b>Selected genes</b>	AADAT,ADSSL1,AGMA,ALDH18A1,ALDH4A1,ALDH5A1,AMDHD1,ARG1,ARG2,ART4,ASL,ASNS,ASNSD1,ASRGL1,ASS1,CAD,CPS1,CTPS1,CTPS2 <sup>#</sup> ,DAO,DDAH1,DDAH2,FAH,FAAH,FAAH2 <sup>#</sup> ,FPGS,FTCD,GAD1,GAD2,GATC,GCLC,GCLM,GFPT1,GFPT2,GGH,GGT1,GLS,GLS2,GLUD1,GLUD2 <sup>#</sup> ,GLUL,GMPS,GOT1,GOT2,HAL,IDH1,LGSN,MECP2 <sup>#</sup> ,MTHFS,NADSYN1,NAGS,NIT2,NOS1,NOS2,NOS3,NOXRED1,OAT,OTC <sup>#</sup> ,PET112,PFAS,PHGDH,PPAT,PRODH,PRODH2,PYCR1,PYCR2,PYCR3,QRSL1,SIRT4,SLC1A3,TAT,UROC1,SLC1A5,SLC6A14 <sup>#</sup> ,SLC6A19,SLC6A15,SLC7A5,SLC7A6,SLC7A8,SLC38A1,SLC38A2,SLC38A3,SLC38A7,SLC38A5 <sup>#</sup> .	<b>84</b> (after removing the 63 duplicated genes)

Keyword: Glutamine

Organism: Homo sapiens

Website: <http://software.broadinstitute.org/gsea/msigdb/search.jsp>

\*(1) Pochini L, Scalise M, Galluccio M, Indiveri C. Membrane transporters for the special amino acid glutamine: structure/function relationships and relevance to human health. Front Chem 2014;2:61.

<sup>#</sup> Genes located in Chromosome X.

**Supporting information Table 2.** Characteristics of the MDACC and Harvard NHS/HPFS studies

Parameter			Frequency		MFT	HR (95% CI) <sup>1</sup>	P <sup>1</sup>
			Patient	Death (%)			
<b>MDACC</b>			858	95 (11.1)	81.1		
Age (years)	≤50		371	31 (8.4)	85.8	1.00	
	>50		487	64 (13.1)	78.1	1.69 (1.10-2.59)	0.017
Sex	Female		362	26 (7.2)	85.9	1.00	
	Male		496	69 (13.9)	77.8	2.07 (1.32-3.25)	0.002
Regional/distantmetastasis	No		709	51 (7.2)	82.7	1.00	
	Yes		149	44 (29.5)	69.4	4.78 (3.19-7.15)	<0.001
Breslow thickness (mm)	≤1		347	7 (2.0)	85.0	1.00	
	>1		511	88 (17.2)	78.1	9.17 (4.25-19.80)	<0.001
Ulceration	No		681	48 (7.1)	84.0	1.00	
	Yes		155	43 (27.7)	64.3	4.91 (3.29-7.42)	<0.001
	Missing		22				
Mitotic rate (mm <sup>2</sup> )	≤1		275	9 (3.3)	82.2	1.00	
	>1		583	86 (14.8)	80.1	4.67 (2.35-9.29)	<0.001
<b>Harvard NHS/HPFS</b>			409	48 (11.5)	179.0		
Age (years)	≤50		72	3 (4.2)	352.5	1.00	
	>50		337	45 (13.4)	167.0	4.04 (1.25–13.06)	0.020
Sex	Female		271	31 (11.4)	198.0	1.00	
	Male		138	17 (12.3)	155.5	1.16 (0.64-2.10)	0.622

Abbreviations: MDACC, The University of Texas MD Anderson Cancer Center; NHS/HPFS: Nurses' Health Study/Health Professionals Follow-up Study; MFT, median follow-up time (months); HR, hazards ratio; CI, confidence interval;

<sup>1</sup>Univariate Cox proportional hazards regression analysis

**Supporting information Table 3.** Comparison of multiple three testing correction methods (FPRP, BFDP and FDR).

<b>SNP</b>	<b><i>P</i></b>	<b>FPRP</b>	<b>BFDP</b>	<b>FDR</b>
rs17676826	0.010	0.081	0.570	0.9538
rs1723518	0.022	0.165	0.292	0.9538
rs9341780	0.022	0.165	0.292	0.9538
rs1007519	0.024	0.179	0.212	0.9538
rs1711934	0.024	0.179	0.253	0.9538
rs12663017	0.024	0.179	0.738	0.9538
rs8012548	0.021	0.169	0.713	0.9538

Abbreviations: SNP, single nucleotide polymorphisms; FPRP, false-positive report probability; BFDP, Bayesian false discovery probability; FDR, false discovery rate.

**Supporting information Table 4.** Function prediction of validated SNPs of genes in the glutamine metabolic pathway

SNP	Gene	Chr	Position (hg 38)	RegDB <sup>1</sup>	Haploreg v4.1 <sup>2</sup>					GRASP QTL hits	Selected eQTL hits	dbSNP func annot
					Promoter histone marks	Enhancer histone marks	DNase	Proteins bound	Motifs changed			
rs17676826	HAL	12	95990567	4		ESC, IPSC, BLD	IPSC		Zbtb3, GCNF, RXRA, Zfp691	1 hit	3 hits	intronic
rs1723518	LGSN	6	63274355	--					MZF1::1-4,NF-kappaB, ZBTB7A		1 hit	
rs9341780	LGSN	6	63274505	5					Myf		2 hits	
rs1007519	LGSN	6	63274970	5					CTCF,GR,Rad21,SMC3		1 hit	
rs1711934	LGSN	6	63275324	--					AP-2, BDP1, SETDB1, Znf143		1 hit	
rs12663017	LGSN	6	63275703	6		GI			Foxa,Foxd1, Foxo-4, HDAC2, Hsf, Maf, TCF12, p300,		3 hits	
rs2253428 <sup>#</sup>	LGSN	6	63276569	--					HNF1,Mef2,TATA		1 hit	3'-UTR
rs2253430 <sup>#</sup>	LGSN	6	63276607	--							1 hit	3'-UTR
rs8012548	NOXRED1	14	77396520	6					Pitx2,Spz1		5 hits	intronic
rs10141317 <sup>\$</sup>	TMED8	14	77377010	4			49 tissues	10 bound proteins	BDP1,Sin3Ak-20, TCF12		3 hits	missense
rs3742737 <sup>\$</sup>	TMED8	14	77377041	2b	24 tissues		52 tissues	12 bound proteins	BHLHE40,CTCF,Smad4		5 hits	missense

Abbreviations: SNP, single nucleotide polymorphisms; Chr, chromosome; dbSNP func annot, dbSNP function annotation; Linkage disequilibrium, LD.

<sup>1</sup>RegulomeDB (<http://www.regulomedb.org>)

<sup>2</sup>HaploReg v4.1 (<http://archive.broadinstitute.org/mammals/haploreg/haploreg.php>)

<sup>#</sup> Functional SNPs have high LD with LGSNrs12663017 ( $r^2=1$ ).

<sup>\$</sup> Functional SNPs have high LD with NOXRED1 rs8012548 (TMED8 rs10141317,  $r^2=0.93$ ; TMED8 rs3742737,  $r^2=0.94$ ).



**Supporting information Table 5.** Stratified analysis of the risk genotypes of selected SNPs in MDACC and Harvard NHS/HPFS studies

Characteristics	0-1 risk genotypes <sup>1</sup>		2-3 risk genotypes <sup>1</sup>		Univariate analysis		Multivariate analysis <sup>2</sup>		<i>P</i> <sub>het</sub> <sup>3</sup>	Interaction <sup>4</sup>
	All	Death (%)	All	Death (%)	HR (95%CI)	<i>P</i>	HR (95%CI)	<i>P</i>		
MDACC										
Age (years)										
≤ 50	185	13 (7.03)	186	18 (9.68)	1.47 (0.72-2.30)	0.292	2.06 (0.97-4.38)	0.059	0.967	0.105
> 50	268	23 (8.58)	219	41 (18.72)	2.27 (1.36-3.79)	0.002	2.02 (1.18-3.45)	0.010		
Sex										
Male	264	26 (9.85)	232	43 (18.53)	1.96 (1.21-3.20)	0.007	1.93 (1.15-3.23)	0.012	0.654	0.229
Female	189	10 (2.29)	173	16 (9.25)	1.85 (0.84-4.07)	0.128	2.42 (1.04-5.63)	0.040		
Regional/distant metastasis										
No	368	15 (4.08)	341	36 (10.56)	2.63 (1.44-4.80)	0.002	2.11 (1.12-3.97)	0.021	0.708	0.521
Yes	85	21 (24.71)	64	23 (35.94)	1.70 (0.94-3.08)	0.079	1.78 (0.95-3.33)	0.071		
Breslow thickness (mm)										
≤1	188	2 (1.06)	159	5 (3.14)	3.03 (0.59-15.63)	0.185	2.81 (0.42-18.81)	0.288	0.802	0.056
>1	265	34 (12.83)	246	54 (21.95)	1.83 (1.19-2.81)	0.006	2.19 (1.36-3.31)	0.001		
Ulceration										
No	363	20 (5.51)	318	28 (8.81)	1.61 (0.91-2.87)	0.102	1.56 (0.87-2.80)	0.135	0.257	0.735
Yes	79	15 (18.99)	76	28 (36.84)	2.26 (1.21-4.23)	0.011	2.60 (1.34-5.04)	0.005		
Missing	22									
Mitotic rate (mm <sup>2</sup> )										
≤1	154	3 (1.95)	121	6 (4.96)	2.67 (0.67-10.70)	0.165	5.64 (0.73-43.80)	0.098	0.316	0.541
>1	299	33 (11.04)	284	53 (18.66)	1.75 (1.13-2.70)	0.012	1.93 (1.23-3.05)	0.005		
NHS/HPFS										
Age (years)										
≤ 50	42	2 (4.76)	30	1 (3.33)	0.71 (0.06-7.80)	0.777	0.67 (0.06-7.40)	0.740	0.422	0.461
> 50	202	21 (10.40)	135	24 (17.78)	1.85 (1.03-3.33)	0.039	1.85 (1.03-3.33)	0.039		
Sex										
Male	86	13 (15.12)	52	4 (7.69)	0.54 (0.17-1.64)	0.274	0.52 (0.17-1.61)	0.260	0.011	0.010
Female	158	10 (6.33)	113	21 (18.58)	3.16 (1.49-6.71)	0.003	3.03 (1.42-6.43)	0.004		

Abbreviations: SNP, single nucleotide polymorphisms; MDACC, The University of Texas MD Anderson Cancer Center; NHS/HPFS: Nurses' Health Study/Health Professionals Follow-up Study; HR, hazards ratio; CI, confidence interval;

<sup>1</sup>Risk genotypes included *HAL* rs17676826 TC+CC, *LGSN* rs12663017 TT and *NOXRED1* rs8012548 AA;

<sup>2</sup>Adjusted for age, sex, Breslow thickness, distant/regional metastasis, ulceration, and mitotic rate in Cox models of SNPs and CMSS in MDACC study and adjusted for age and sex in Harvard study;

<sup>3</sup> $P_{\text{het}}$ , heterogeneity test between subgroups.

<sup>4</sup>Interaction: The interaction between the risk genotypes and each clinical variable.

## Reference

1. Pochini L, Scalise M, Galluccio M, Indiveri C. Membrane transporters for the special amino acid glutamine: structure/function relationships and relevance to human health. *Front Chem* 2014;2:61.

